

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号

特開平5-307569

(43)公開日 平成5年(1993)11月19日

(51)Int.Cl.⁵

G 0 6 F 15/40

識別記号

5 0 0 T

庁内整理番号

7060-5L

F I

技術表示箇所

審査請求 未請求 請求項の数4(全 9 頁)

(21)出願番号 特願平4-112682
(22)出願日 平成4年(1992)5月1日

(71)出願人 000004226
日本電信電話株式会社
東京都千代田区内幸町一丁目1番6号
(72)発明者 小澤 英昭
東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内
(72)発明者 中川 透
東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内
(74)代理人 弁理士 伊東 忠彦

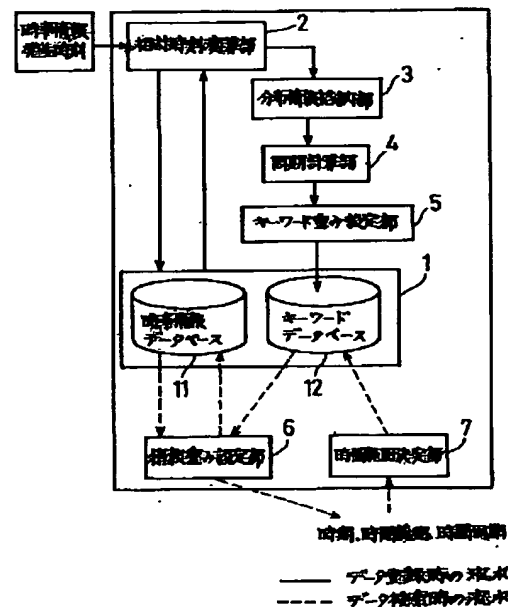
(54)【発明の名称】 時間変動する情報に対応する情報の蓄積及び検索方法

(57)【要約】

【目的】 本発明の目的は、ユーザが要求する時間周期に最適な情報をデータベースより提供することである。

【構成】 本発明は、類似した情報の発生回数を予め指定された時間範囲に区切って計測する相対時刻演算部2と、複数の時間範囲内の類似した情報の発生頻度のばらつきを計測して格納する周期計算部3と、発生頻度のばらつきの情報から情報発生する周期を調べる周期計算部4、各情報に対する発生の周期を含めた情報を記憶するデータベース1と、ユーザから与えられた時間周期に対する各情報の発生頻度から時間周期に対して重要な情報を選択する情報重み設定部6から構成される。

本発明の一実施例の情報蓄積、検索方法のシステム構成図



【特許請求の範囲】

【請求項1】 周期的に現れる第1の情報と該第1の情報の特徴を表現する一つ以上の第2の情報と該第1の情報が発生した時刻が格納されるデータベースにおいて、類似した第1の情報の発生回数を予め指定された時間範囲に区切って計測し、

複数の該時間範囲内の類似した第1の情報の発生頻度のばらつきを計測し、

該発生頻度のばらつきの情報から第2の情報が発生する周期を調べ、

各第2の情報に対する発生の周期を含めた第3の情報を記憶し、

ユーザから与えられた時間周期に対する第2の情報の発生の頻度から時間周期に対して重要度を決定し、重要度の高い第1の情報を選択することを特徴とする時間変動する情報に対応する情報の蓄積及び検索方法。

【請求項2】 周期的に現れる第1の情報と該第1の情報の特徴を表現する一つ以上の第2の情報と該第1の情報が発生した時刻が格納されるデータベースにおいて、前記第2の情報が発生する時間周期を計算するために、前記第1の情報に1つ以上のキーワードを与え、時間周期に対する各キーワードの発生分布を計測し、

予め与えられる計算式により、該キーワードの時間周期に対する発生の頻度より該キーワードを有する第1の情報の発生する時間周期を決定することを特徴とする請求項1記載の時間変動する情報に対応する情報の蓄積及び検索方法。

【請求項3】 周期的に現れる第1の情報と該第1の情報の特徴を表現する一つ以上の第2の情報と該第1の情報が発生した時刻が格納されるデータベースにおいて、前記キーワードの時間周期に対する発生頻度の分布と予め与えられている時間周期に対する発生頻度の分布とを比較して前記キーワードを持つ情報の時間周期に対する重要度の計算への影響を変化させることを特徴とする請求項1記載の時間変動する情報に対応する情報の蓄積及び検索方法。

【請求項4】 周期的に現れる第1の情報と該第1の情報の特徴を表現する一つ以上の第2の情報と該第1の情報が発生した時刻が格納されるデータベースにおいて、前記キーワードと共に、時間周期の情報と、時間周期に対する重みの情報を格納するデータベースを備え、予め与えられた時間範囲と時間周期に対してキーワードの重みの順に並べて予め与えられる閾値と重みのデータを比較し、

該当する時間周期と重みを大きい順に順序付けして並べ、該時間範囲と時間周期に対して重要なキーワードを判定することを特徴とする請求項1記載の時間変動する情報に対応する情報の検索方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、時間変動する情報に対応する情報の蓄積・検索方法に係り、特に、情報の格納、検索を目的とするデータベースシステムにおいて、周期的に重要になる情報と突発的な情報と分離または選択を行う情報の格納、検索を行う時間変動する情報に対応する情報の蓄積・検索方法に関する。

【0002】

【従来の技術】情報検索システムでの検索、特に時事情報を検索する場合には、獲得したい情報の時間的な領域を制限するために検索条件の一つとして、例えば、“1990年1月”のように時間範囲の条件を与えることができる。図書館のような紙を媒体とするシステムでも製本された雑誌のバックナンバーから特定の年の雑誌を選択するような方法により、情報全体に一律な時間区間の制約を与えて検索を行う方法がある。

【0003】このように、従来は同時に発生した情報を時間軸に対して同一の価値があるものとして扱う。

【0004】

【発明が解決しようとする課題】しかしながら、時事情報の価値は例えば、終戦記念日の情報や台風の情報のように毎年発生する情報や、ハレー彗星のように数十年に一度の周期で話題になるような、時間に対して周期的に価値を持つ情報もあれば、情報が発生した時点では有効であるが、すぐに価値を失う突発的な情報等様々であり、全ての情報に対して同じ時間区間で、情報の廃棄や検索の制限を行うことはできない。ところが従来のシステムでは、同時に発生した情報を時間軸に対して同一の価値があるものとして扱うために、記憶媒体の容量の制限等から、単純に新しい情報を高速な記憶媒体に保存したり、単純に時間が経過した情報を除去してしまい、季節の話題のような予め情報が必要となる時期を予測できる情報を的確に入手しにくいという問題がある。

【0005】本発明は上記の点に鑑みなされたもので、ユーザが要求する時間周期に最適な情報を提供することができる時間変動する情報に対応する情報の蓄積・検索方法を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明は、周期的に現れる第1の情報と該第1の情報の特徴を表現する一つ以上の第2の情報と第1の情報が発生した時刻が格納されるデータベースにおいて、類似した第1の情報の発生回数を予め指定された時間範囲に区切って計測し、複数の時間範囲内の類似した第1の情報の発生頻度のばらつきを計測し、発生頻度のばらつきの情報から第2の情報が発生する周期を調べ、各第2の情報に対する発生の周期を含めた第3の情報を記憶し、ユーザから与えられた時間周期に対する第2の情報の発生の頻度から時間周期に対して重要度を決定し、重要度の高い第1の情報を選択する。

【0007】また、本発明は周期的に現れる第1の情報と第1の情報の特徴を表現する一つ以上の第2の情報と第1の情報が発生した時刻が格納されるデータベースにおいて、第2の情報が発生する時間周期を計算するために、第1の情報に1つ以上のキーワードを与え、時間周期に対する各キーワードの発生分布を計測し、予め与えられる計算式により、キーワードの時間周期に対する発生頻度よりキーワードを有する第1の情報の発生する時間周期を決定する。

【0008】また、本発明は周期的に現れる第1の情報と第1の情報の特徴を表現する一つ以上の第2の情報と第1の情報が発生した時刻が格納されるデータベースにおいて、キーワードの時間周期に対する発生頻度の分布と予め与えられている時間周期に対する発生頻度の分布とを比較してキーワードを持つ情報の時間周期に対する重要度の計算への影響を変化させる。

【0009】さらに、本発明は周期的に現れる第1の情報と第1の情報の特徴を表現する一つ以上の第2の情報と第1の情報が発生した時刻が格納されるデータベースにおいて、キーワードと共に、時間周期の情報と、時間周期に対する重みの情報を格納するデータベースを備え、予め与えられた時間範囲と時間周期に対してキーワードの重みの順に並べて予め与えられる閾値と重みのデータを比較し、該当する時間周期と重みを大きい順に順序付けして並べ、時間範囲と時間周期に対して重要なキーワードを判定する。

【0010】

【作用】本発明は、まず、1日、1週間、1年等のように予め指定された時間範囲に区切って、類似した情報の発生頻度を計測し、発生頻度のばらつきの情報から情報が発生する周期を調べ、各情報と、情報の発生周期を含めた情報をデータベースに記憶し、ユーザから与えられた時間周期に対する各情報の発生頻度より時間周期に対して重要な情報を選択する。

【0011】また、情報の時間周期を計算するために、情報にキーワードを与え、時間周期に対する各キーワードの発生分布を計測し、所定の計算式により、個々のキーワードの時間周期に対する発生頻度より、そのキーワードを持つ情報が発生する時間周期を決定する。

【0012】また、時間周期に対してキーワードの発生頻度がランダムな場合には、類似の計算に対しての影響を小さくし、キーワードの発生する時間周期の特徴が顕著なキーワードに対して類似度の計算への影響を大きくする。

【0013】さらに、キーワード、時間周期の情報及び時間周期に対する重みの情報を格納するデータベースを用いて、特に所定のある時間範囲と、時間周期、例えば、3月から4月の時間範囲で1年周期に対して、キーワードの重みの順に並べて予め与えられる閾値と重みのデータを比較することにより該当する時間周期と時間範

囲における重要なキーワードを判定できる。

【0014】このように、本発明は、情報の発生頻度により情報を区別することで、原爆の日のような周期的に話題になる情報と、突発的な事件のような、一過性の情報とを分離して扱えることと、8月から9月に周期的に話題となるキーワードは台風であるといった、任意の時間範囲と時間周期に対応して重要となるキーワードを識別することを最も主要な特徴であり、従来の技術では検索対象の期間の全ての情報は同等に検索されていたが、これにより、検索対象の時期にあった情報が的確に検索される。

【0015】

【実施例】本発明の実施例を図面を用いて説明する。

【0016】図1は本発明の一実施例の情報蓄積、検索のシステム構成を示す。

【0017】本発明のシステムは、時事情報データベース11とキーワードデータベース12からなるデータベース1と、相対時刻演算部2、分布情報格納部3、周期計算部4、キーワード重み設定部5、情報重み設定部6、時間範囲決定部7より構成される。

【0018】データベース1は、時事情報を格納するための時事情報データベース11とキーワード及び各キーワードの相対時刻に対する重みと、検索用の絶対時刻を格納するキーワードデータベース12の7つのモジュールより構成される。

【0019】本実施例において、時事情報の構造は一つ以上の内容に関するデータと、少なくとも一つ以上のキーワード、情報の発生時間を持つ。情報の発生時間は、例えば各時事情報が新聞に掲載されたり、テレビやラジオにより放送された日付を用いることで実現できる。

【0020】相対時刻演算部2は、データベース1から検索された情報と入力された情報の発生時刻をデータベース1が格納する最も古いデータから相対的な時刻に変換する。

【0021】分布情報格納部3は、相対時刻演算部2から出力された相対時刻と各キーワードによる2次元的なスロットをもち、相対時刻演算部2で決められた時刻とキーワードから該当するスロットの値をインクリメントし、保存する。

【0022】周期計算部4は、分布情報格納部3に格納された値の分布から情報の発生する周期を判定する。

【0023】キーワード重み設定部5は、周期計算部4からの周期の情報に基づいて、各キーワードの周期に対する重みを計算する。

【0024】情報重み設定部6は、データベース1の検索時に、キーワード毎に付けられた重みの情報から各時事情報毎の重みを計算する。

【0025】時間範囲決定部7は、検索する対象となる時間周期の幅を計算する。

【0026】図1において、実線は、データベース1へ

のデータ登録時の流れを示し、点線は、データベース1のデータ検索時の流れを示す。

【0027】データベース1に投入される各時事情報は、情報を識別するために、ジャンルや情報の特徴を表す複数個のキーワードと、その情報が発生した時を示す発生時間が付けられている。

【0028】本実施例における分布情報格納部3は、例えば、時間範囲として、1日から31日、32日から61日のように、一月毎に時間範囲を与える。そして、相対時刻演算部2で与えられた時刻は、365日で除して剰余を求め、閏年分を換算して3月1日を第1日として1年の範囲内の相対的な日数に変換し、該当するスロットの値を増加させる。この結果、例えば、年末の宝くじの情報は、276日から306日までのスロットの所に値が集中するので、年末の宝くじの情報は、276日から306日の範囲である12月に重要な情報として判断できる。

【0029】図2は、本発明の一実施例の時事情報の登録時の処理過程のフローチャートを示す。

【0030】ステップ20：データベース1に新しい時事情報が追加される場合には、時事情報と発生時間を入力する。

【0031】ステップ21：新しい時事情報の持つ各キーワードを用いて、時事情報データベース11を検索し、該当するキーワードを持つ情報の発生時間を獲得する。

【0032】ステップ22：時事情報データベース11より検索された発生時間を相対時刻演算部2によって予め定められた時刻から相対時間に変換し、キーワードとともに、分布情報格納部3に格納する。

【0033】ステップ23からステップ26の処理をキーワードの個数分だけ繰り返す。

【0034】ステップ23：分布情報格納部3よりキーワードを1つ取り出す。

【0035】ステップ24：周期計算部4において分布情報格納部3からある一つのキーワードに対する分布のデータからピークの部分の時間範囲を求める。

【0036】ステップ25：キーワード重み設定部5において、各キーワード毎の分布の値を予め設定した閾値のパターンと比較する。

【0037】ステップ26：キーワードの時間周期に対する分布のうち、ピーク値の閾値よりも大きい場合には、その値を周期の値及び位相差の値と共に、キーワードデータベース12に格納する。

【0038】キーワードデータベース12に格納されるデータの例としては、例えば「台風」というキーワードがあったとすると、上記の一実施例による分布情報格納部3を用いれば、「台風」と8月に相当する位相差の値として154日から184日と周期の値として1年に相当する365日、8月の重みの値と9月についても同様

にして得られた値が格納される。この過程を入力すべき時事情報（キーワード）が存在する限り繰り返す。

【0039】また、この処理過程は、データベース中のデータが全くない場合でも、同様であり、十分なデータが格納されると安定する。

【0040】本実施例におけるキーワード重み設定部5での処理として、一月刻みで設定された時間範囲のスロット値から、平均値を求め、平均値2倍未満の時間範囲に対する重みはゼロ、平均値の2倍以上の値に対する重みは1であり、3倍以上の値に対する重みは2のように、平均値のn倍以上の値の重みをn-1とする。さらに、この各時間範囲の重みの値をその時間範囲に対するスロットの値を掛けて、そのキーワードに対する重みとする。

【0041】この結果、例えば台風というキーワードは、3月や4月には、1回程度しか現れないが、6月過ぎからしばしば登場するようになり、8月には25回位、9月には40回位現れる。よって平均すると、月に8回程度登場するため、8月は重みが $25/8 \times 25$ で75となり、9月は200となる。

【0042】これに対して株価というキーワードは、月単位でみると、1年を通しておよそ、40から60程度で登場するため、全ての月で重みが0になり、時間周期に依存した情報ではないと判断される。

【0043】本実施例で示した時事情報データベース11を利用するユーザがある季節によく起こる時事情報を検索したい場合、例えば、ユーザは8月、1ヵ月、1年周期などの検索したい時期や日付と時間範囲と時間周期の実施例のシステムに投入する。

【0044】図3は本発明の一実施例のデータベースを利用するユーザからの検索要求に対する処理の過程を示すフローチャートである。図1においては、点線で示される流れを参照されたい。

【0045】ステップ30：時期、時間範囲、時間周期が入力される。

【0046】ステップ31：時間範囲決定部7は、ステップ30で与えられた日時と時間範囲から、例えば、1年周期という条件から350日から380日間という検索すべき時間周期の幅に変換する。これは、入力された時間周期を日を単位とした周期に変換している。

【0047】ステップ32：ステップ31で決定した時間の幅と8月、1ヵ月間を表す、154日から184日という時間範囲を条件としてキーワードデータベース12を検索し、該当するキーワードと、キーワードの重みを獲得する。これらは、一年周期で8月に重要となるキーワードなので、以下では8月のキーワードと呼ぶ。

【0048】ステップ33：キーワードデータベース12を検索した結果からキーワードを1つ取り出し、以下の処理がすべての検索結果の処理が終了するまで、繰り返す。

【0049】ステップ34：キーワードデータベース12から検索された個々の情報が持つキーワードを抽出し、そのキーワードを用いて、時事情報データベース11を検索する。

【0050】ステップ35：情報重み設定部6において、例えば、ステップ34で抽出された時事情報のキーワードと、ステップ32で得られた8月のキーワードを比較し、一致するキーワードについては、キーワード重みを付け、一致しないキーワードについては0として、個々の時事情報の8月に対する重みの値の総和を求め

る。

【0051】ステップ36：すべての検索結果の処理を終了したら、情報重み設定部6で作成された重みの大きい順に順序付けして並べ、ユーザに提供する。

【0052】図4は本発明の一実施例におけるキーワードデータベースのデータ構造を示す。キーワードデータベース12の構造としては、1つのキーワードに対して、少なくとも1つ以上の周期情報からなる。周期情報は、本発明の方法によって得られた1年等の時間周期の情報と9月などの位相差の情報、時間周期に対する重みの情報から構成される。同図の例では、「台風」というキーワードは、350日～380日の周期で、3月1日から見て154日から184日の位相（8月に相当）を持ち、重みが75の周期情報と、350日～380日周期で、3月1日から見て185日から214日の位相（9月に相当）を持ち、重みが200の周期情報から構成される。

【0053】次に本発明の他の実施例について説明する。

【0054】分布情報格納部3は時間範囲のスロットをデータベース1中に格納している情報のうち最も古い情報の時刻から、一日刻みの相対的な時刻の範囲として表現し、周期計算部4は分布情報格納部3で得られた分布の自己相関関数を求め、時間遅れを計算することにより、月単位のような固定的な時間範囲ではなく、任意の日数を持つ周期を求めることができる。

【0055】上記の他の実施例では、自己相関関数を使用した、任意の周期を求めるために、フーリエ変換を用いることも可能であり、パワースペクトルを求めることで、情報が発生する時間周期と発生頻度を計算できる。さらに、フーリエ変換を用いることで、周期の情報と共に、位相差の情報を得ることができる。

【0056】次に、キーワード重み設定部5の他の実施例について説明する。各時間範囲のスロットの値の対数関数 \log_2 を求め、予め与えた閾値以上の場合には、対数関数 \log_2 の値を重みとし、閾値以下の場合には重みを0とする（但し、スロットの値が0の場合にはも、重みを0とする）。この結果、重みの大きい領域では \log 関数によって重みの大きさが均一化されるために、ある特定のキーワードの値による影響が大きくなり過ぎることが

なくなる。

【0057】さらにキーワード重み設定部5の他の実施例として、フーリエ変換を用いてキーワードの発生頻度を周波数領域に変換した場合において、パワースペクトルが予め与えられた閾値よりも大きい場合には、その値をキーワードの重みとし、閾値よりも小さい場合には0とする。そして、位相差の情報から相対的な日数として時間周期の情報を獲得する。

【0058】次に、データベース検索時の他の実施例について説明する。先に示した図3のステップ32において検索したキーワードを該当する時間範囲の重みの情報を用いて値の大きい方から順序付けし、予め与えられている閾値よりも大きな重みを持つ情報をユーザに提供することにより、ある時間範囲における重要なキーワードを得ることも可能である。

【0059】例えば、図3のステップ32におけるデータベースの検索によって、8月のキーワードとして「台風」（重みの値75）、「夏」（重みの値100）、「水不足」（重みの値50）、「ビール」（重みの値30）等が得られたとすると、これらのキーワードに付けられている重みのデータの順に並べ、例えば、閾値として50を与えたならば、「夏」、「台風」、「水不足」のように8月に重要なキーワードを選択して得ることができる。

【0060】ユーザが与える時間周期の値の範囲は、一実施例では最小は1日、最大で時事情報データベース11に格納している最も古い情報の発生日時から現在までの日数である。周期の指定の単位は、日数だけでなく一週間、一年間などの指定も可能である。この場合、指定された周期を標準的な日数の周期に変換し、予め指定された幅を持たすことで達成できる。例えば、一年周期であれば、標準的な365日に変換し、さらに、前後例えば5%の幅を持たせ、347日から383日として、検索を行う。一実施例での時間周期の単位は一日であるが、時事情報の発生時間のサンプリングを変えることで、任意の値にすることができる。

【0061】上記のように、本発明による方法は、任意の時間範囲と時間周期に対応して重量なキーワードを識別することができる。このキーワードでデータベースに蓄積されている情報より適正な情報を得ることができる。

【0062】

【発明の効果】上述のように本発明によれば、例えば、株価の情報や、交通事故の情報のような、季節や時間の周期に無関係に発生する一過性の情報は、時間周期に対する重みが小さくなる。これに対して、お祭りなどの情報は情報発生時点から時間が経つにつれて情報の重みは小さくなるが、毎年繰り返して関連する情報が発生するために、一年周期の時間範囲で見ると、他の情報に比べて情報の重みが大きくなり、周期的に重要な情報と日

常的に発生する情報や突発的な情報を分離することが可能となる。

【0063】その結果、日々キーワード付きの情報を本発明の方法を用いて、データベースに登録していくことにより、ユーザが要求する時間周期に最適な情報を提供することができる。

【図面の簡単な説明】

【図1】本発明の一実施例の情報蓄積・検索方法のシステム構成図である。

【図2】本発明の一実施例の時事情報の登録時の処理過程のフローチャートである。

【図3】本発明の一実施例のデータベースを利用するユーザからの検索要求に対する処理の過程を示すフローチ

ャートである。

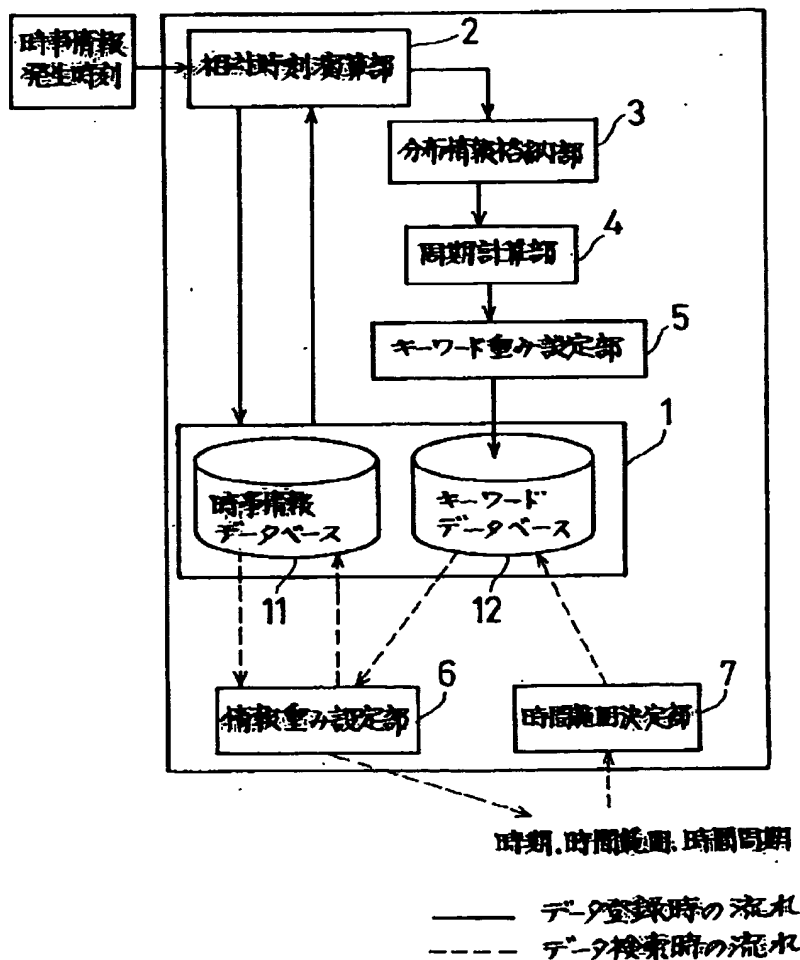
【図4】本発明の一実施例におけるキーワードデータベースのデータ構造を示す図である。

【符号の説明】

- 1 データベース
- 2 相対時刻演算部
- 3 分布情報格納部
- 4 周期計算部
- 5 キーワード重み設定部
- 6 情報重み設定部
- 7 時間範囲決定部
- 11 時事情報データベース
- 12 キーワードデータベース

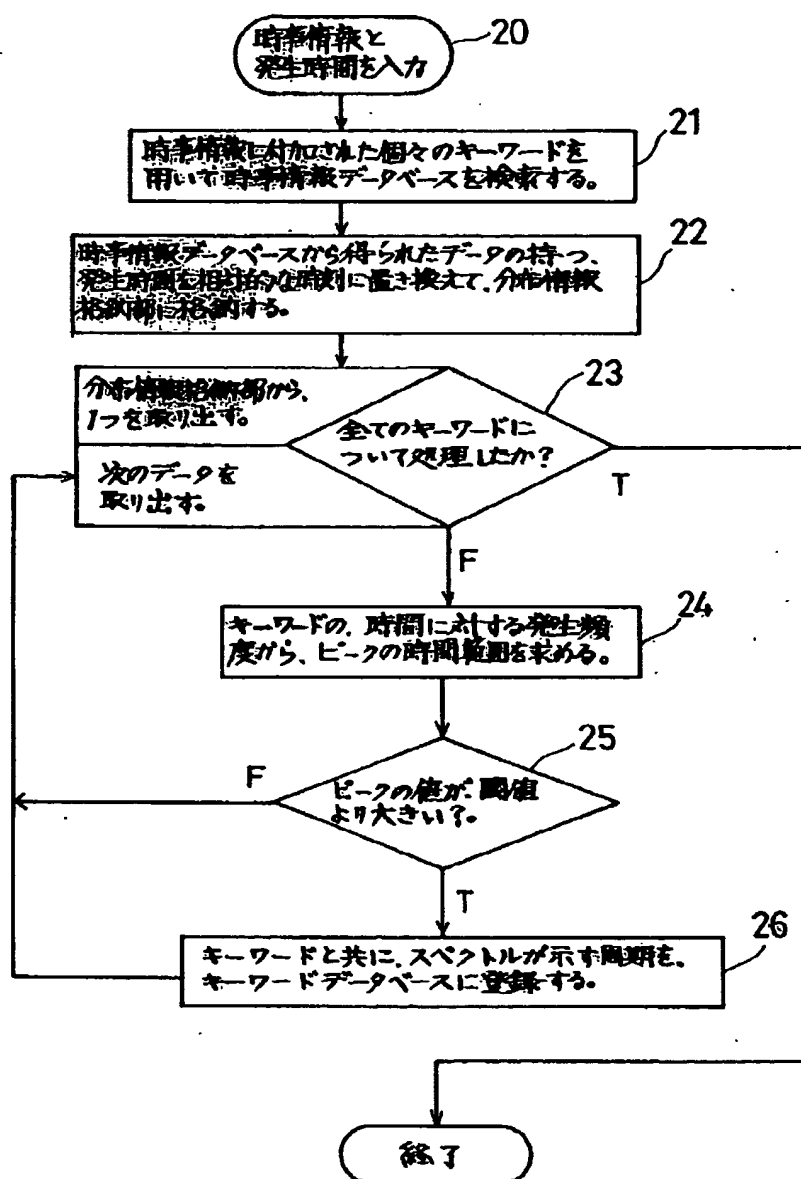
【図1】

本発明の一実施例の情報蓄積・検索方法のシステム構成図



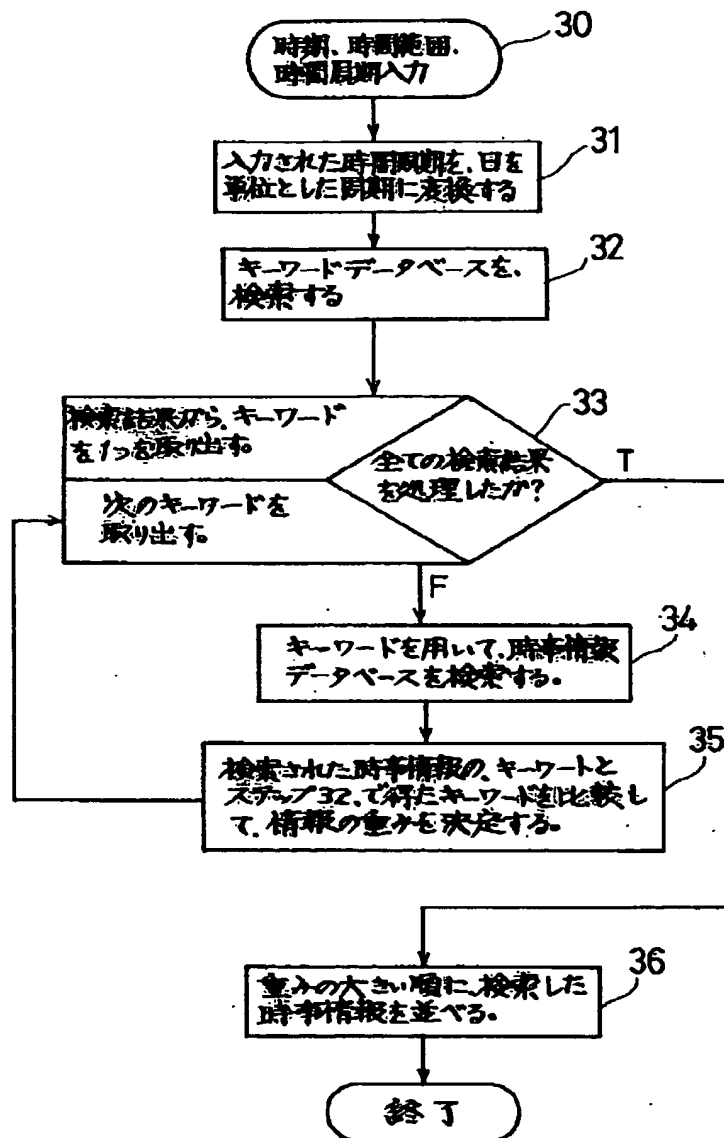
【図2】

本発明の一実施例の新規の時事情報の登録時の
処理過程のフローチャート



【図3】

本発明の一実施例のデータベースを利用するユーザからの検索要求に対する処理の過程を示すフローチャート



本発明の一実施例におけるキーワードデータベースのデータ構造を示す図

[illegible]

PTO 03-5742

Japanese Patent

Document No. 5-307569

**ACCUMULATION AND SEARCH METHOD FOR INFORMATION CORRESPONDING TO
INFORMATION THAT VARIES OVER TIME**

[Jikan Hendo Suru Joho ni Taio Suru Joho no Chikuseki Oyobi

Kensaku Hoho]

Hideaki Ozawa and Toru Nakagawa

UNITED STATES PATENT AND TRADEMARK OFFICE

Washington, D.C.

October 2003

Translated by: Schreiber Translations, Inc.

Country : Japan

Document No. : 5-307569

Document Type : Kokai

Language : Japanese

Inventor : Hideaki Ozawa and Toru Nakagawa

Applicant : Nihon Telegraph & Telephone Co.,
Ltd.

IPC : G 06 F 15/40

Application Date : May 1, 1992

Publication Date : November 19, 1993

Foreign Language Title : Jikan Hendo Suru Joho ni Taio Suru
Joho no Chikuseki Oyobi Kensaku
Hoho

English Title : ACCUMULATION AND SEARCH METHOD FOR
INFORMATION CORRESPONDING TO
INFORMATION THAT VARIES OVER TIME

(54) Title of the invention

An accumulation and search method for information corresponding to information that varies over time

(57) Summary

Objective: The objective of the present invention is to provide, from a database, information optimal for a temporal periodic frequency requested by a user.

Constitution: The constitution of the present invention includes the relative time computation unit (2), which measures the generation frequencies of analogous sets of information as they are compartmentalized in preliminarily designated time frames, the periodic frequency calculation unit [sic: Presumably "distribution information storage unit," see below] (3), which measures and stores the irregularities of the generation frequencies of analogous sets of information within multiple time frames, the periodic frequency calculation unit (4), which investigates, from the sets of information on the irregularities of the generation frequencies, the information generation periods, the database (1), which memorizes sets of information inclusive of the generation periods of the respective sets of information, and the information weight designation unit (6), which selects, based on the generation frequencies of the respective sets of information on a temporal periodic frequency requested by a user,

¹ Numbers in the margin indicate pagination in the foreign text.

sets of information important for said temporal periodic frequency.

Patent Claims

/2

Claim 1

An accumulation and search method for information corresponding to information that varies over time with the following characteristics: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information,

The generation frequencies of analogous sets of first information as they are compartmentalized in preliminarily designated time frames are measured, whereas

The irregularities of the generation frequencies of the analogous sets of first information within a multiple number of said time frames are measured, whereas

The generation periods of the second information are investigated based on the information on said generation frequency irregularities, whereas

A set of third information inclusive of the generation frequencies of the respective sets of second information is memorized, whereas

The level of significance for a temporal periodic frequency requested by a user is determined based on the generation frequency of the second information for said temporal periodic frequency, whereas

A set of first information with a high level of significance is selected.

Claim 2

An accumulation and search method for information corresponding to information that varies over time specified in Claim 1 with the following characteristics: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information,

At least one key word is assigned to the aforementioned first information for the purpose of calculating the generation temporal periodic frequency for said second information, whereas

The generation distribution of each key word with respect to [said] temporal periodic frequency is measured, whereas

The generation temporal periodic frequency for the first information which bears said key word is determined, by using a preliminarily orchestrated calculation formula, from the generation frequency of said key word with respect to said temporal periodic frequency.

Claim 3

An accumulation and search method for information corresponding to information that varies over time specified in Claim 1 with the following characteristics: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information,

The generation frequency distribution of the aforementioned key word with respect to each temporal periodic frequency is compared with a preliminarily assigned generation frequency distribution with respect to said temporal periodic frequency for the purpose of varying the effect exerted on the calculation of the level of significance for said temporal periodic frequency by the information which bears the aforementioned key word.

Claim 4

An accumulation and search method for information corresponding to information that varies over time specified in Claim 1 with the following characteristics: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information,

A database which stores not only the aforementioned key word but also sets of information on temporal periodic frequencies, and sets of weight information on said temporal periodic frequencies is configured, whereas

The respective key words are listed in the order of weight in relation to the preliminarily assigned time frame and temporal periodic frequency for the purpose of comparing preliminarily assigned threshold values and weight data, whereas the corresponding temporal periodic frequencies and weights are ranked and listed in the order of magnitude for the purpose of judging key words important for said time frames and temporal periodic frequencies.

Detailed explanation of the invention

[0001]

(Industrial application fields)

The present invention concerns an accumulation and search method for information corresponding to information that varies over time, and in particular, it concerns an accumulation and search method for information corresponding to information that varies over time with the following characteristics: In a database system designed to serve information storage and search functions, sets of information which become periodically important and sets of serendipitously arising information are separated or selected in the context of information storage and search.

[0002]

(Prior art)

In a case where a search, especially a current event information search, is launched by using an information search system, a time frame condition such as "January 1990" may, for example, be stipulated as one of the search conditions for limiting the temporal regions for the targeted information. In the cases of systems which use paper media such as libraries, too, there exists a known method wherein uniform time frame restrictions are imposed on the entire sets of information and wherein a search is launched based on a format whereby magazines of particular years are selected from the back numbers of bound volumes.

[0003]

Thus, simultaneously arising sets of information are assumed to bear identical values in relation to the time axis in the prior art.

[0004]

The values of sets of information on current events, however, are diverse, for certain sets bear periodic values over time, as in the cases of information arising each year (e.g., information on end of war memorial days, information on typhoons, etc.) and information which becomes popular at a frequency of once every several decades (e.g., Halley's comet, etc.), whereas other sets of information are serendipitous in that, although they may be discussed often initially, they quickly lose their values over time, and thus, it is impossible to impose information discard and search restrictions on all sets of information by using identical

time frames. Simultaneously arising sets of information, however, are handled by the system of the prior art based on the assumption that they bear identical values in relation to the time axis, and since the limitations on the memory medium capacities, etc. mandate that the latest sets of information be routinely saved into high-speed memory media and that outdated sets of information be routinely deleted, it is difficult to ensure the acquisition of sets of information that predict in advance the time zones over which the information is vital (e.g., seasonal topics, etc.), which is problematic.

[0005]

The objective of the present invention, which has been conceived in acknowledgment of the aforementioned backdrop, is to provide an accumulation and search method for information corresponding to information that varies over time in the context of providing information optimal for a temporal periodic frequency requested by a user.

[0006]

(Mechanism for solving the problems)

The following constitution is provided by the present invention: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information, the generation frequencies of analogous sets of first information

as they are compartmentalized in preliminarily designated time frames are measured, whereas the irregularities of the generation frequencies of the analogous sets of first information within a multiple number of said time frames are measured, whereas the generation periods of the second information are investigated based on the information on said generation frequency irregularities, whereas a set of third information inclusive of the generation frequencies of the respective sets of second information is memorized, whereas the level of significance for a temporal periodic frequency requested by a user is determined based on the generation frequency of the second information for said temporal periodic frequency, whereas a set of first information with a high level of significance is selected.

[0007]

/3

The following constitution may, furthermore, be provided by the present invention: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information, at least one key word is assigned to the aforementioned first information for the purpose of calculating the generation temporal periodic frequency for said second information, whereas the generation distribution of each key word with respect to [said] temporal periodic frequency is measured, whereas the generation temporal periodic frequency for the first information which bears

said key word is determined, by using a preliminarily orchestrated calculation formula, from the generation frequency of said key word with respect to said temporal periodic frequency.

[0008]

The following constitution may, furthermore, be provided by the present invention: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information, the generation frequency distribution of the aforementioned key word with respect to each temporal periodic frequency is compared with a preliminarily assigned generation frequency distribution with respect to said temporal periodic frequency for the purpose of varying the effect exerted on the calculation of the level of significance for said temporal periodic frequency by the information which bears the aforementioned key word.

[0009]

The following constitution may, furthermore, be provided by the present invention: In a database constituted to store a set of first information which appears periodically, at least one set of second information which expresses the attributes of said first information, and the generation time of said first information, a database which stores not only the aforementioned key word but also sets of information on temporal periodic frequencies, and sets of weight information on said temporal periodic frequencies is configured, whereas the respective key words are listed in the

order of weight in relation to the preliminarily assigned time frame and temporal periodic frequency for the purpose of comparing preliminarily assigned threshold values and weight data, whereas the corresponding temporal periodic frequencies and weights are ranked and listed in the order of magnitude for the purpose of judging key words important for said time frames and temporal periodic frequencies.

[0010]

(Functions)

As far as the present invention is concerned, the time is initially compartmentalized into preliminarily designated time frames (e.g., 1 day, 1 week, 1 year, etc.), and after the generation frequencies of analogous sets of information have been measured, the information generation periods are investigated based on information on the irregularities of the generation frequencies, and after data inclusive of the respective sets of information and information generation periods have been memorized into the database, sets of information important for a temporal periodic frequency requested by a user are selected based on the generation frequencies of the respective sets of information for said temporal periodic frequency.

[0011]

In order to calculate the information temporal periodic frequency, furthermore, a key word(s) is assigned to information, and after the generation frequency of each key word in relation to

each temporal periodic frequency has been measured, the generation temporal periodic frequency for the information that bears said key word(s) is determined, by using a certain calculation formula, based on the generation frequencies of the individual key words with respect to said temporal periodic frequency.

[0012]

In a case where the key word generation frequency of the key word(s) is random in relation to the temporal periodic frequency, furthermore, the effect on the analog calculation is minimized, whereas the effect on the calculation of the analogous degree of the key word characterized by a prominent feature of the key word generation temporal periodic frequency is maximized [sic: Non sequitur].

[0013]

It is also possible to use a database designed to store key words, information on the temporal periodic frequency, and information on the weight for each temporal periodic frequency, according to which the key words are listed in the order of weight in relation to a certain time frame and a certain temporal periodic frequency (e.g., 1 year periodic frequency over a time frame from March to April), whereas preliminarily assigned threshold values are compared with the weight data for the purpose of judging key words important for the corresponding temporal periodic frequency and time frame.

[0014]

Thus, as far as the present invention is concerned, sets of information are mutually distinguished in terms of the information generation frequency, based on which periodically popularized information (e.g., Atomic Bomb Day, etc.) can be handled separately from transient information (e.g., unexpected events, etc.), whereas its most notable characteristic lies in its ability to identify key words which are important in relation to arbitrary time frames and temporal periodic frequencies (e.g., "typhoon" as a periodically popularized key word between August and September), and in contradistinction with the embodiment of the prior art, wherein all sets of information over a targeted search time frame are equally searched for, the [selective?] search for information over the targeted search time frame can be ensured by this constitution.

[0015]

(Application examples)

Application examples of the present invention will be explained with reference to figures.

[0016]

Figure 1 shows the constitution of the information accumulation and search system of an application example of the present invention.

[0017]

The system of the present invention is constituted not only by the database (1) inclusive of the current event information

database (11) and the key word database (12) but also by the relative time computation unit (2), the distribution information storage unit (3), the periodic frequency calculation unit (4), the key word weight designation unit (5), the information weight designation unit (6), and the time frame designation unit (7).

[0018]

The database (1) is characterized by a 7-module constitution inclusive of the current event information database (11) for storing current event information and the key word database (12) for storing key words, weights of the respective key words in relation to the relative time, and the absolute time for search.

[0019]

The current event information is constituted to include at least one content datum, at least one key word, and the information generation time. The information generation time may, for example, be coded by the dates of the listings of the respective sets of current event information on newspapers or of their television or radio broadcasts.

[0020]

The relative time computation unit (2) converts the respective generation times of the information searched from the database (1) and the inputted information into relative times in relation to the oldest data stored in the database (1).

[0021]

The distribution information storage unit (3) is constituted not only to possess two-dimensional slots assigned to the relative

times outputted from the relative time computation unit (2) and to the respective key words but also to increment and save the values of the corresponding slots from the times and key words determined by the relative time computation unit (2).

[0022]

The periodic frequency calculation unit (4) judges the information generation temporal periodic frequency based on the distribution of values stored in the distribution information storage unit (3).

[0023]

The key word weight designation unit (5) calculates the weights of the respective key words in relation to the temporal periodic frequency based on the temporal periodic frequency information obtained from the periodic frequency calculation unit (4).

[0024]

The information weight designation unit (6) calculates, during the search of the database (1), the weights of the respective sets of information on current events based on the sets of weight information assigned to the respective key words.

[0025]

The time frame designation unit (7) calculates the expanse of the temporal periodic frequency targeted for the search.

[0026]

In Figure 1, the solid line shows the flow of an operation for registering data into the database (1), whereas the dotted line shows the flow of an operation for the data search of the

/4

database (1).

[0027]

In order to enable information identification, the respective sets of current event information encoded into the database (1) are designated to bear multiple key words which express genres and information attributes as well as the times of the raisings of said sets of information.

[0028]

The distribution information storage unit (3) of the present application example is constituted to assign a time frame corresponding to a month (e.g., day 1 through day 31, day 32 through day 61, etc.). The time rendered by the relative time computation unit (2), furthermore, is divided by 365 (days) for calculating the remainder in consideration of leap years and then converted into a relative date over a 1-year period whereby March 1 is designated as day 1, and the value of the corresponding slot is incremented. As a result, the values of the sets of information on end-of-the-year lotteries can be acknowledged to be concentrated within the slot from day 276 to day 306, based on which the information on the end-of-the-year lotteries can be acknowledged to be important information for December, which coincides with the range of day 276 through day 306.

[0029]

Figure 2 shows a flow chart for a routine for registering the current event information in [said] application example of the present invention.

[0030]

Step 20: In a case where a new set of current event information becomes added to the database (1), the current event information and its generation time are inputted.

[0031]

Step 21: The current event information database (11) is searched by using the respective key words specific to the new set of current event information, as a result of which the generation time for the information that bears said key words becomes ascertained.

[0032]

Step 22: The generation time searched from the current event information database (11) is converted from the preliminarily designated time into the relative time by the relative time computation unit (2), and the latter is then stored in the distribution information storage unit (3) together with the key words.

[0033]

The respective routines from the step 23 to the step 26 are executed over a repetition frequency corresponding to the number of key words.

[0034]

Step 23: One key word is retrieved from the distribution information storage unit (3).

[0035]

Step 24: The peak portion time frame is calculated by the periodic frequency calculation unit (4) from data on the distribution of said single key word retrieved from the distribution information storage unit (3).

[0036]

Step 25: The distribution value specific to each key word is compared with a preliminarily designated threshold value pattern by the key word weight designation unit (5).

[0037]

Step 26: Of the distributions of the respective key words in relation to the temporal periodic frequency, ones the peak values of which exceed the threshold value are identified, and such values are stored in the key word database (12) together with the temporal periodic frequency value and the phase differential value.

[0038]

As far as examples of data stored in the key word database (12) are concerned, in a case where "typhoon" is designated as a key word, for example, the duration from day 154 to day 184 becomes selected by using the distribution information storage unit (3) of the aforementioned application example as the phase differential value corresponding to "typhoon," whereas a period of 365 days, which is equivalent to 1 year, becomes designated as the

temporal periodic frequency value, whereas the weight value specific to August and similarly ascertained values for September become stored. The foregoing processes are repeated until all the inputted sets of current event information (i.e., key words) have become exhausted.

[0039]

This routine process is likewise applicable to a case where absolutely no data are present in the database, and a stable state becomes achieved upon the storage of sufficient data.

[0040]

As far as the routine executed by the key word weight designation unit (5) of the present application example is concerned, the average value is calculated from the slot value of a time frame designated for each month, and "0" is designated as the weight specific to a time frame shorter than a double of the average value, whereas "1" in a case where the corresponding weight is at least a double of the average value, whereas "2" in a case where the corresponding weight is at least a triple of the average value; thus, $n-1$ is designated as the weight value of a case where the prevailing value is higher than a value obtained by multiplying the average value by n . The weight value of each time frame, furthermore, is multiplied by the slot value corresponding to said time frame, and the obtained value is designated as the weight specific to the specified key word.

[0041]

As a result, the key word "typhoon" appears only once or so in March or April, but it begins to appear often after June, and its appearance frequencies in August and September are approximately 25 times and 40 times, respectively. The average appearance frequency is therefore approximately 8 times per month, and accordingly, the weighted value for August is $25/8 \times 25 = 75$, whereas the corresponding value for September is 200.

[0042]

The key word "stock price," in contrast, consistently appears at frequencies of approximately 40 to 60 times per month over an entire year, and since the weight is 0 for every month, this information is judged not to depend on the temporal periodic frequency.

[0043]

In a case where a user who accesses the current event information database (11) shown in the present application example wishes to search for information on an event which occurs frequently during a given season, for example, the user encodes such data as the search periods, dates, temporal durations, and the temporal periodic frequency (e.g., August, one month, 1-year period, etc.) into the system of the application example.

[0044]

Figure 3 is a flow chart which shows the flow of routines executed in response to a search request issued from the user who accesses the database of [said] application example of the present

invention. The flow expressed by the dotted line in Figure 1 should be referred to.

[0045]

Step 30: The period, time frame, and the temporal periodic frequency are inputted.

[0046]

Step 31: The time frame designation unit (7) may, for example, change the search condition of 1-year period into the time span from Day 350 to day 380 based on the dates and time frame ascertained at the step 30. It is thus that the inputted temporal periodic frequency is converted into a period constituted by day units.

[0047]

Step 32: A key word database (12) search is launched by stipulating the condition determined at the step 31, namely the time span from day 154 to day 184, from which "August" and "1-month period" can be inferred, followed by the acquisitions of corresponding key words and key word weights. These are key words which become important in August at a 1-year periodic frequency, and therefore, they will hereafter be referred to as "August key words."

[0048]

Step 33: A single key word is retrieved based on the results of the key word database (12) search, and the following routines are repeated until all the search results have become processed.

[0049]

/5

Step 34: Key words which the individual sets of information obtained as a result of the key word database (12) search bear are extracted, and a current event information database (11) search is launched by using said key words.

[0050]

Step 35: The current event information extracted at the step 34 and the August key words obtained at the step 32 may, for example, be compared by the information weight designation unit (6), and coinciding key words are weighted, whereas "0" is assigned to non-coinciding key words, followed by the calculation of the total sum of the weight values for the individual sets of current event information with regard to August.

[0051]

Step 36: After all the search results have become processed, they are ranked in the order of weights formulated by the information weight designation unit (6) and then provided for the user.

[0052]

Figure 4 shows the data structure of the key word database of [said] application example of the present invention. The structure of the key word database (12) is characterized by the assignment of at least one set of temporal periodic frequency information to a single key word. The sets of temporal periodic frequency information are constituted by the data ascertained by

the method of the present invention, namely the temporal periodic frequency information (e.g., 1 year, etc.), phase differential information (e.g., September, etc.), and the weight for the temporal periodic frequency. In the example of the same figure, the key word "typhoon" is constituted not only by a set of temporal periodic frequency information which coincides with the phase from day 154 to day 184 from the standpoint of March 1 at a periodic frequency of 350 days ~ 380 days (equivalent to August) and which is weighted at 75 but also by a set of temporal periodic frequency information which coincides with the phase from day 185 to day 214 from the standpoint of March 1 at a periodic frequency of 350 days ~ 380 days (equivalent to September) and which is weighted at 200.

[0053]

Next, another application example of the present invention will be explained.

[0054]

Of the information data on the time frame slots stored into the database (1), the distribution information storage unit (3) expresses [the data] in terms of a relative time range incremented daily from the time of the oldest information, whereas the periodic frequency calculation unit (4) calculates the autocorrelation of the distribution obtained from the distribution information storage unit (3) and then computes the time delay, based on which it becomes possible to calculate a period

constituted by an arbitrary number of days rather than a fixed time frame (e.g., month unit, etc.).

[0055]

The autocorrelation was used in the aforementioned alternative application example, although it is also possible to rely on a Fourier conversion in the context of calculating an arbitrary period, whereas the time period over which information becomes generated and the generation frequency can be computed by calculating the power spectrum. Not only the period information but also phase differential information can, furthermore, be obtained by relying on the Fourier conversion.

[0056]

Next, another application example of the key word weight designation unit (5) will be explained. The relative function \log_2 of the slot value of each time frame is calculated, and in a case where it exceeds a preliminarily designated threshold value, the value of the relative function \log_2 is designated as a weight, whereas in a case where it is lower than the threshold value, "0" is designated as a weight (in a case where the slot value is "0," too, "0" is designated as a weight). As a result, the weight magnitude becomes homogenized by the log function within a heavily weighted region, and accordingly, excessively high effects of the values of certain key words can be avoided.

[0057]

As far as still another application example of the key word weight designation unit (5) is concerned, furthermore, in a case

where the key word generation frequency has been compartmentalized into a frequency region by means of Fourier conversion and where the concomitantly obtained power spectrum exceeds a preliminarily designated threshold value, the prevailing value is designated as a key word weight, whereas "0" is designated in a case where the same is lower than the threshold value. The time period information is then acquired, as a relative number of days, from the phase differential information.

[0058]

Next, another application example pertaining to the database search will be explained. The key words obtained as a result of the search launched at the step 32 of Figure 3 shown earlier are ranked in the order of value magnitudes by using the corresponding time frame weight information, and key words important during a certain time frame can be obtained by providing the user with sets of information which bear weights heavier than the preliminarily designated threshold value.

[0059]

In a case where "typhoon" (weight value: 75), "summer" (weight value: 100), "water shortage" (weight value: 50), "beer" (weight value: 30), etc. are assumed to have been obtained as a result of the database search launched at the step 32 in Figure 3, for example, the data are ranked in the order of the weights assigned to the respective key words, and in a case where "50" is designated as the threshold value, for example, key words

important for August such as "summer," "typhoon," and "water shortage" can be selected.

[0060]

According to one application example, the minimal time period range assigned by the user is 1 day, whereas the maximal range coincides with the time span from the generation date of the oldest information stored in the current event information database (11) to the present date. Not only the number of days but 1-week and 1-year periods can also be designated as period designation units. The goal of the latter case can be achieved by converting the designated period into the standard period constituted by the number of days and by then assigning a preliminarily designated margin to the obtained value. In a case where a one-year period is designated, for example, the duration is converted into standard 365 days, whereas 5% margins are added respectively to the beginning and end of said period, and a search is launched with regard to the period from day 347 to day 383. The time period unit in [said] application example is one day, although an arbitrary value can be designated by varying the sampled current event information generation time.

[0061]

As has been mentioned above, the method of the present invention enables the identifications of important key words corresponding to arbitrary time frames and temporal periodic frequencies. Optimal information can be obtained from the sets of

information accumulated in the database under the pervasion of such key words.

[0062]

(Effects of the invention)

As has been mentioned above, as far as the present invention is concerned, sets of transient information which arise regardless of seasons or time periods (e.g., stock price information, traffic accident information, etc.) are weighted lightly in relation to the temporal periodic frequency. In contrast, the weights of such sets of information as festival information, etc. become lighter over time since the time of the information generation, but since related sets of information become repeatedly generated every year, the weight of such information is heavier than those of other sets of information with regard to the time frame of a one-year period, and thus, it becomes possible to separate periodically important information from information which becomes generated daily and information which becomes generated unexpectedly.

/6

[0063]

As a result, it becomes possible to provide the information optimal for the temporal periodic frequency requested by a user by registering, into a database on a daily basis, sets of information which bear key words by using the method of the present invention.

Brief explanation of the figures

Figure 1: A systems constitutional diagram pertaining to the information accumulation and search method of an application example of the present invention.

Figure 2: A flow chart pertaining to a sequence of routines executed in the context of registering current event information in [said] application example of the present invention.

Figure 3: A flow chart which shows the sequence of routines executed in response to a search routine request submitted by a user who accesses the database of [said] application example of the present invention.

Figure 4: A diagram which shows the data structure of the key word database of [said] application example of the present invention.

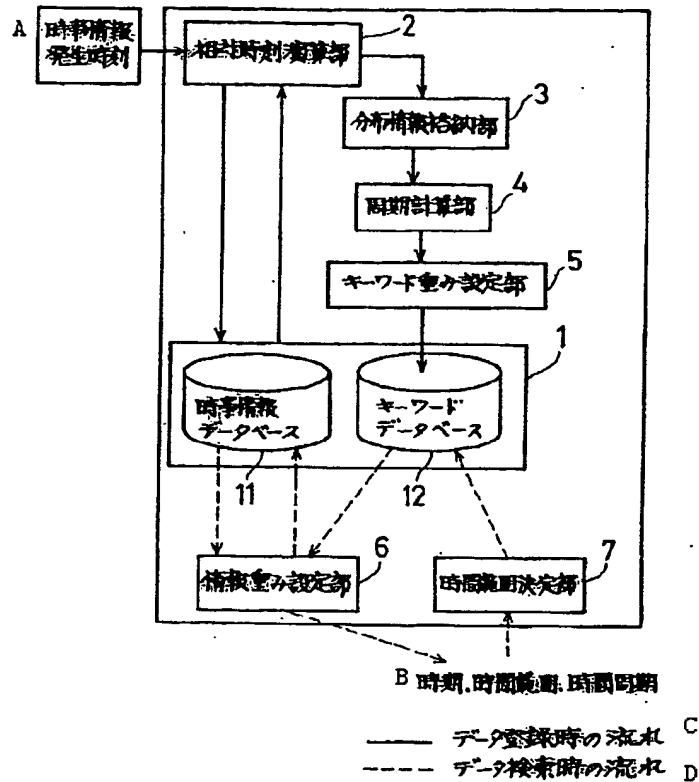
(Explanation of notations)

- (1): Database;
- (2): Relative time computation unit;
- (3): Distribution information storage unit;
- (4): Periodic frequency calculation unit;
- (5): Key word weight designation unit;
- (6): Information weight designation unit;
- (7): Time frame designation unit;
- (11): Current event information database;
- (12): Key word database.

Figure 1

【図1】

○ 本発明の一実施例の情報蓄積・検索方法のシステム構成図



[(0): Systems constitutional diagram pertaining to the information accumulation and search method of an application example of the present invention; (A): Current event information generation time; (B): Time period, time frame, and temporal periodic frequency; (C): Flow at the time of data registration; (D): Flow at the time of data search; (2): Relative time computation unit; (3): Distribution information storage unit; (4): Periodic frequency calculation unit; (5): Key word weight designation unit; (6): Information weight designation unit; (7): Time frame designation

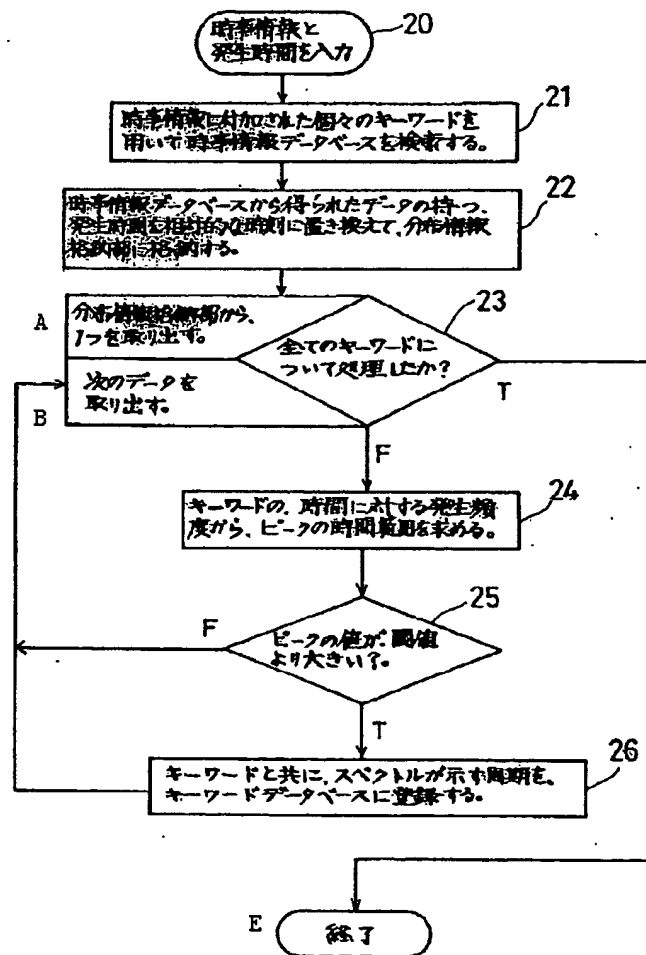
unit; (11): Current event information database; (12): Key word database]

Figure 2

/7

【図2】

本発明の一実施例の新規の時事情報の登録時の処理過程のフローチャート



[(0): Flow chart pertaining to a sequence of routines executed in the context of registering new current event information in an

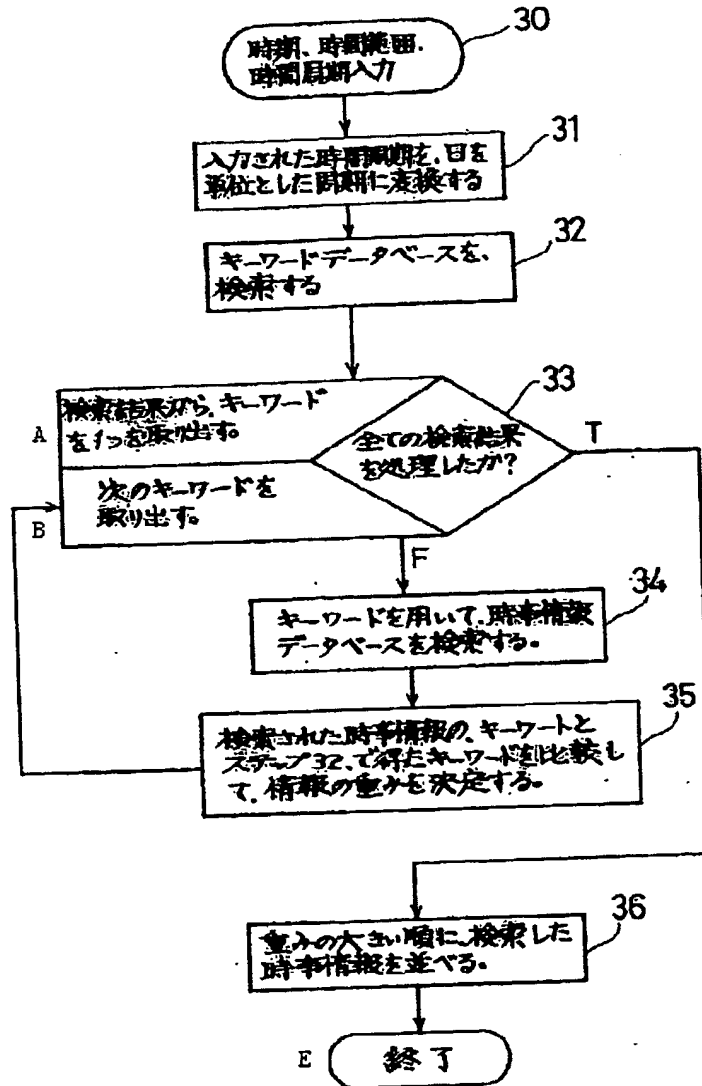
application example of the present invention; (A): One key word is retrieved from the distribution information storage unit; (B): Next data are retrieved; (E): End; (20): The current event information and generation time are inputted; (21): The current event information database is searched by using the individual key words added to the current event information; (22): The generation times possessed by the data obtained from the current event information database are converted into relative times and then stored in the distribution information storage unit; (23): Have all key words been processed?; (24): The peak time frame is calculated based on the generation frequencies of the key words per unit time; (25): Are peak values higher than the threshold value?; (26): The periodic frequency indicated by the spectrum is registered into the key word database together with the key words]

Figure 3

/8

【図3】

本発明の一実施例のデータベースを利用するユーザからの検索要求に対する処理の過程を示すフローチャート



[(0): Flow chart which shows the sequence of routines executed in response to a search routine submitted by a user who accesses the database of an application example of the present invention; (A):

One key word is retrieved from the search results; (B): Next key word is retrieved; (E): End; (30): The period, time frame, and temporal periodic frequency are inputted; (31): The inputted temporal periodic frequency is converted into a period expressed by the day unit; (32): The key word database is searched; (33): Have all the search results been processed?; (34): The current event information database is searched by using the key words; (35): The information weights are determined by comparing the key words of the current event information obtained as a result of the search with the key words obtained at the step 32; (36): The sets of current event information obtained as a result of the search are ranked in the order of weight]

Figure 4

19

①	キーワード	⑦ 前期1			⑦ 前期2			⑦ 前期3			
		前期	位相	重み	前期	位相	重み	前期	位相	重み	
②	台風	350~ 380	154~ 184	75	350~ 380	185~ 214	200		185~ 214	30	...
③	水不足	350~ 380	123~ 153	80	350~ 380	154~ 184	120				...
④	ビール	350~ 380	62~ 72	20	350~ 380	154~ 184	30	350~ 380	185~ 214	30	...
⑤	終戦記念日	350~ 380	154~ 184	200							...
⑥	決算	350~ 380	62~ 72	300	350~ 380	236~ 275	40				...

本発明の一実施例におけるキーワードデータベースのデータ構造を示す図

【図4】

[(0): Diagram which shows the data structure of the key word database of an application example of the present invention; (1): Key word; (2): Typhoon; (3): Water shortage; (4): Beer; (5): End of war memorial day; (6): Settlement of accounts; (7): Period; (8): Phase; (9): Weight]